MAS223 Statistical Modelling and Inference

Exercises

The exercises are grouped into sections, corresponding to chapters of the lecture notes. Within each section exercises are divided into warm-up questions, ordinary questions, and challenge questions. Note that there are no exercises accompanying Chapter 8.

The vast majority of exercises are ordinary questions. Ordinary questions will be used in homeworks and tutorials; they cover the material content of the course. Warm-up questions are typically easier, often nothing more than revision of relevant material from first year courses. Challenge questions are typically harder and test ingenuity.

Contents

1 Univariate Distribution Theory

Warm-up Questions

1.1 Let X be a random variable taking values in $\{1, 2, 3\}$, with $\mathbb{P}[X = 1] = \mathbb{P}[X = 2] = 0.4$. Find $\mathbb{P}[X = 3]$, and calculate both $\mathbb{E}[X]$ and $\mathrm{Var}[X]$.

1.2 Let Y be a random variable with probability density function (p.d.f.) $f(y)$ given by

$$f(y) = \begin{cases} y/2 & \text{for } 0 \le y < 2; \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability that Y is between $\frac{1}{2}$ and 1. Calculate $\mathbb{E}[Y]$ and $\mathrm{Var}[Y]$.

Ordinary Questions

1.3 Define $F : \mathbb{R} \to [0, 1]$ by

$$F(y) = \begin{cases} 0 & \text{for } y \le 0; \\ y^2 & \text{for } y \in (0, 1); \\ 1 & \text{for } y \ge 1. \end{cases}$$

(a) Sketch the function F, and check that it is a distribution function.

(b) If Y is a random variable with distribution function F, calculate the p.d.f. of Y.

1.4 Let X be a discrete random variable, taking values in $\{0, 1, 2\}$, where $\mathbb{P}[X = n] = \frac{1}{3}$ for $n \in \{0, 1, 2\}$. Sketch the distribution function $F_X : \mathbb{R} \to \mathbb{R}$.

1.5 Define $f : \mathbb{R} \to [0, 1]$ by

$$f(x) = \begin{cases} 0 & \text{for } x < 0; \\ e^{-x} & \text{for } x \ge 0. \end{cases}$$

(a) Show that f is a probability density function.

(b) Find the corresponding distribution function and evaluate $\mathbb{P}[1 < X < 2]$.

1.6 Sketch graphs of each of the following two functions, and explain why each of them is not a distribution function.

(a) $F(x) = \begin{cases} 0 & \text{for } x \le 0; \\ x & \text{for } x > 0. \end{cases}$

(b) $F(x) = \begin{cases} 0 & \text{for } x < 0; \\ x + \frac{1}{4}\sin 2\pi x & \text{for } 0 \le x < 1; \\ 1 & \text{for } x \ge 1. \end{cases}$

1.7 Let $k \in \mathbb{R}$ and define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} k(x - x^2) & \text{for } x \in (0,1), \\ 0 & \text{otherwise.} \end{cases}$$

Find the value of k for which $f(x)$ is a probability density function, and calculate the probability that X is greater than $\frac{1}{2}$.

1.8 The probability density function $f(x)$ is given by

$$f(x) = \begin{cases} 1 + x & \text{for } -1 \leq x < 0; \\ 1 - x & \text{for } 0 \leq x < 1; \\ 0 & \text{otherwise.} \end{cases}$$

Find the corresponding distribution function $F(x)$ for all real x.

1.9 Let

$$F(x) = \frac{e^x}{1 + e^x} \quad \text{for all real } x.$$

(a) Show that F is a distribution function, and find the corresponding p.d.f. f.

(b) Show that $f(-x) = f(x)$.

(c) If X is a random variable with this distribution, evaluate $\mathbb{P}[|X| > 2]$.

1.10 The discrete random variable X has the probability function

$$\mathbb{P}[X = x] = \begin{cases} \frac{1}{x(x+1)} & \text{for } x \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases}$$

(a) Use the partial fractions of $\frac{1}{x(x+1)}$ to show that $\mathbb{P}[X \leq x] = 1 - \frac{1}{x+1}$, for all $x \in \mathbb{N}$.

(b) Write down the distribution function $F(x)$ of X, for $x \in \mathbb{R}$. Sketch its graph. What are the values of $F(2)$ and $F(\frac{3}{2})$?

(c) Evaluate $\mathbb{P}[10 \leq X \leq 20]$.

(d) Is $\mathbb{E}[X]$ defined? If so, what is $\mathbb{E}[X]$? If not, why not?

Challenge Questions

1.11 For which values of $r \in [0, \infty)$ is $\int_1^\infty x^{-r}\, dx$ finite? Give an example of a continuous random variable for which $\mathbb{E}[X]$ is defined but $\mathbb{E}[X^2]$ is not.

1.12 Let $A = \mathbb{Q} \cap [0,1]$. Since A is countable, let us write $A = (a_1, a_2, a_3, \ldots)$, where $a_n \neq a_m$ for $n \neq m$. Let $F : \mathbb{R} \to \mathbb{R}$ be given by

$$F(x) = \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}_{\{a_n \leq x\}}.$$

Here, $\mathbb{1}_{\{a_n \leq x\}}$ is equal to 1 if $a_n \leq x$, and equal to 0 otherwise.

(a) Show that F is a distribution function.

(b) Let X be a random variable with distribution function F. Show that $\mathbb{P}[X = x] = 0$ for all $x \in \mathbb{R} \setminus A$, and that $\mathbb{P}[X = a_n] = 2^{-n}$ for all $x \in A$.

2 Standard Univariate Distributions

Warm-up Questions

2.1 (a) A standard fair dice is rolled 5 times. Let X be the number of sixes rolled. Which standard distribution (and which parameters) would you use to model X?

(b) A fair coin is flipped until the first head is shown. Let X be the total number of flips, including the final flip on which the first head appears. Which standard distribution (and which parameter) would you use to model X?

Ordinary Questions

2.2 Let $\lambda > 0$. Write down the p.d.f. f of the random variable X, where $X \sim Exp(\lambda)$, and calculate its distribution function F. Hence, show that $\frac{f(t)}{1-F(t)}$ is constant for $t > 0$.

2.3 Let $\lambda > 0$ and let X be a random variable with $Exp(\lambda)$ distribution. Let $Z = \lfloor X \rfloor$, that is let Z be X rounded down to the nearest integer. Show that Z is geometrically distributed with parameter $p = 1 - e^{-\lambda}$.

2.4 Let $\mu \in \mathbb{R}$. Let X_1 and X_2 be independent random variables with distributions $N(\mu, 1)$ and $N(\mu, 4)$, respectively. Let T_1, T_2 and T_3 be defined by

$$T_1 = \frac{X_1 + X_2}{2}, \qquad T_2 = 2X_1 - X_2, \qquad T_3 = \frac{4X_1 + X_2}{5}.$$

Find the mean and variance of T_1, T_2 and T_3. Which of $\mathbb{E}[T_1]$, $\mathbb{E}[T_2]$ and $\mathbb{E}[T_3]$ would you prefer to use as an estimator of μ?

2.5 Let X be a random variable with $Ga(\alpha, \beta)$ distribution.

(a) Let $k \in \mathbb{N}$. Show that

$$\mathbb{E}[X^k] = \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{\beta^k}.$$

Hence, calculate $\mu = \mathbb{E}[X]$ and $\sigma^2 = \mathrm{Var}(X)$ and verify that these formulas match the ones given in lectures.

(b) Show that $\mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{2}{\sqrt{\alpha}}$.

2.6 (a) Using `R`, you can obtain a plot of, for example, the p.d.f. of a $Ga(3, 2)$ random variable between 0 and 10 with the command

```
curve(dgamma(x,shape=3,scale=2),from=0,to=10)
```

Use `R` to investigate how the shape of the p.d.f. of a Gamma distribution varies with the different parameter values. In particular, fix a value of β, see how the shape changes as you vary α.

(b) Investigate the effect that changing parameters values has on the shape of the p.d.f. of the Beta distribution. To produce, for example, a plot of the p.d.f. of $Be(4, 5)$, use

```
curve(dbeta(x,shape1=4,shape2=5),from=-1,to=2)
```

2.7 Suggest which standard discrete distributions (or combination of them) we should use to model the following situations.

(a) Organisms, independently, possess a given characteristic with probability p. A sample of k organisms with the characteristic is required. How many organisms will need to be tested to achieve this sample?

(b) In Texas Hold'em Poker, players make the best hand they can by combining two cards in their hand with five 'community' cards that are placed face up on the table. At the start of the game, a player can only see their own hand. The community cards are then turned over, one by one.

A player has two hearts in her hand. Three of the community cards have been turned over, and only one of them is a heart. How many hearts will appear in the remaining two community cards?

Find the probability of seeing $k = 0, 1, 2$ hearts.

2.8 Let X be a $N(0, 1)$ random variable. Use integration by parts to show that $\mathbb{E}[X^{n+2}] = \frac{1}{n+1}\mathbb{E}[X^n]$ for any $n = 0, 1, 2, \ldots$. Hence, show that

$$\mathbb{E}[X^n] = \begin{cases} 0 & \text{if } n \text{ is odd;} \\ \frac{1}{(1)(3)(5)\ldots(n-1)} & \text{if } n \text{ is even.} \end{cases}$$

2.9 Let $X \sim N(\mu, \sigma^2)$. Show that $\mathbb{E}[e^X] = e^{\mu + \frac{\sigma^2}{2}}$.

Challenge Questions

2.10 Let X be a random variable with a continuous distribution, and a strictly increasing distribution function F. Show that $F(X)$ has a uniform distribution on $(0, 1)$.

Suggest how we might use this result to simulate samples from standard distributions.

2.11 Prove that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

3 Transformations of Univariate Random Variables

Warm-up Questions

3.1 Let $f : \mathbb{R} \to [0, \infty)$ by $f(x) = (x-1)^2$. Identify the subinterval of $x \in \mathbb{R}$ for which $f(x) = (x-1)^2$ is strictly increasing. Find an inverse function to f on this subinterval and identify the set of x for which $f(x) \leq 1$.

Ordinary Questions

3.2 Let $\alpha, \beta > 0$ and let X be a random variable with the $Be(\alpha, \beta)$ distribution. Show that $Y = 1 - X$ has the $Be(\beta, \alpha)$ distribution.

3.3 Let X be a random variable with the uniform distribution on $(0, 1)$, and let $\lambda > 0$. Show that $Y = \frac{-\log(X)}{\lambda}$ has an $Exp(\lambda)$ distribution.

3.4 Let $\alpha > 0$ and let $X \sim Be(\alpha, 1)$ distribution. Let $Y = \sqrt[r]{X}$ for some positive integer r. Show that Y also has a Beta distribution, and find its parameters.

3.5 Let X have a uniform distribution on $[-1, 1]$. Find the p.d.f. of $|X|$ and identify $|X|$ as a standard distribution.

3.6 Let $\alpha, \beta > 0$ and let $X \sim Be(\alpha, \beta)$. Let $c > 0$ and define $Y = c/X$.

(a) Suppose that $\alpha > 1$. Show that

$$\mathbb{E}[Y] = \frac{cB(\alpha - 1, \beta)}{B(\alpha, \beta)} = \frac{c(\alpha + \beta - 1)}{\alpha - 1}.$$

(b) Find the p.d.f. of Y.

*If $\alpha \in [0, 1]$ then $\mathbb{E}[Y]$ is not defined, see **3.12**.*

3.7 Let Θ be an angle chosen according to a uniform distribution on $(-\frac{\pi}{2}, \frac{\pi}{2})$, and let $X = \tan \Theta$. Show that X has the Cauchy distribution.

3.8 Let X be a random variable with the p.d.f.

$$f(x) = \begin{cases} 1 + x & \text{for } -1 < x < 0; \\ 1 - x & \text{for } 0 < x < 1; \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability density functions of

(a) $Y = 5X + 3$

(b) $Z = |X|$

3.9 Let X have the uniform distribution on $[a, b]$.

(a) For $[a, b] = [-1, 1]$, find the p.d.f. of $Y = X^2$.

(b) For $[a,b] = [-1,2]$, find the p.d.f. of $Y = |X|$.

3.10 Let X have a uniform distribution on $[-1,1]$ and define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(x) = \begin{cases} 0 & \text{for } x \leq 0; \\ x^2 & \text{for } x > 0. \end{cases}$$

Find the distribution function of $g(X)$.

3.11 Let X be a random variable with the Cauchy distribution. Show that X^{-1} also has the Cauchy distribution.

Challenge Questions

3.12 In Question **3.6**, if $\alpha \leq 1$ show that $\mathbb{E}[Y]$ is not defined.

3.13 If we were to pretend that $g(x) = 1/x$ was strictly monotone, we could (incorrectly) apply Lemma 3.1 and use the formula $f_Y(y) = f_X(g^{-1}(y))|\frac{dg^{-1}}{dy}|$ to solve **3.11**. We would still arrive at the correct answer. Can you explain why?

4 Multivariate Distribution Theory

Warm-up questions

4.1 Let $T = \{(x, y) \,:\, 0 < x < y\}$. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x,y) = \begin{cases} e^{-2x-y} & \text{for } (x,y) \in T; \\ 0 & \text{otherwise.} \end{cases}$$

Sketch the region T. Calculate $\int_0^\infty \int_0^\infty f(x,y)\,dx\,dy$ and $\int_0^\infty \int_0^\infty f(x,y)\,dy\,dx$, and verify that they are equal.

4.2 Sketch the following regions of $\mathbb{R}^2$.

(a) $S = \{(x, y) \,:\, x \in [0, 1],\ y \in [0, 1]\}$.

(b) $T = \{(x, y) \,:\, 0 < x < y\}$.

(c) $U = \{(x, y) \,:\, x \in [0, 1],\ y \in [0, 1],\ 2y > x\}$.

Ordinary Questions

4.3 Let (X, Y) be a random vector with joint probability density function

$$f_{X,Y}(x,y) = \begin{cases} ke^{-(x+y)} & \text{if } 0 < y < x \\ 0 & \text{otherwise.} \end{cases}$$

(a) Using that $\mathbb{P}[(X, Y) \in \mathbb{R}^2] = 1$, find the value of k.

(b) For each of the regions S, T, U in **4.2**, calculate the probability that (X, Y) is inside the given region.

(c) Find the marginal p.d.f. of Y, and hence describe the distribution of Y as a standard distribution.

4.4 Let $S = [0, 1] \times [0, 1]$, and let U and V have joint probability density function

$$f_{U,V}(u,v) = \begin{cases} \frac{4u+2v}{3} & (u,v) \in S; \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $\mathbb{P}[U + V \leq 1]$.

(b) Find $\mathbb{P}[V \leq U^2]$.

4.5 For the random variables U and V in Exercise **4.4**:

(a) Find the marginal p.d.f. $f_U(u)$ of U.

(b) Find the marginal p.d.f. $f_V(v)$ of V.

(c) For v such that $f_V(v) > 0$, find the conditional p.d.f. $f_{U|V=v}(u)$ of U given $V = v$.

(d) Check that each of f_U, f_V and $f_{U|V=v}$ integrate over $\mathbb{R}$ to 1.

(e) Calculate the two forms of conditional expectation, $\mathbb{E}[U|V=v]$, and $\mathbb{E}[U|V]$.

4.6 Let (X,Y) be a random vector with joint probability density function

$$f_{X,Y}(x,y)=\begin{cases}\frac{y-x}{2} & x\in[-1,0], y\in[0,1];\\ \frac{x+y}{2} & x\in[0,1], y\in[0,1];\\ 0 & \text{otherwise.}\end{cases}$$

Find the marginal p.d.f. of X. Show that the correlation coefficient X and Y is zero. Show also that X and Y are not independent.

4.7 Let X and Y be independent random variables, with $0<\mathrm{Var}(X)=\mathrm{Var}(Y)<\infty$. Let $U=X+Y$ and $V=XY$. Show that U and V are uncorrelated if and only if $\mathbb{E}[X]+\mathbb{E}[Y]=0$.

4.8 Let $\lambda>0$. Let X have an $Exp(\lambda)$ distribution, and conditionally given $X=x$ let U have a uniform distribution on $[0,x]$. Calculate $\mathbb{E}[U]$ and $\mathrm{Var}(U)$.

4.9 Let $k\in\mathbb{R}$ and let (X,Y) have joint probability density function

$$f_{X,Y}(x,y)=\begin{cases}kx\sin(xy) & \text{for } x\in(0,1), y\in(0,\pi),\\ 0 & \text{otherwise.}\end{cases}$$

(a) Find the value of k.

(b) For $x\in(0,1)$, find the conditional probability density function of Y given $X=x$.

(c) Find $\mathbb{E}[Y|X]$.

4.10 Let U have a uniform distribution on $(0,1)$, and conditionally given $U=u$ let X have a uniform distribution on $(0,u)$.

(a) Find the joint p.d.f of (X,U) and the marginal p.d.f. of X.

(b) Show that $\mathbb{E}[U|X=x]=\frac{x-1}{\log x}$.

4.11 Let (X,Y) have a bivariate distribution with joint p.d.f. $f_{X,Y}(x,y)$. Let $y_0\in\mathbb{R}$ be such that $f_Y(y_0)>0$. Show that $f_{X|Y=y_0}(x)$ is a probability density function.

Challenge Questions

4.12 Give an example of random variables (X,Y,Z) such that

$$\min\big\{\mathbb{P}[X<Y],\mathbb{P}[Y<Z],\mathbb{P}[Z<X]\big\}=\frac{2}{3}.$$

5 Transformations of Multivariate Distributions

Warm-up Questions

5.1 (a) Define $u = x$ and $v = 2y$. Sketch the images of the regions S, T and U from question **4.2** in the (u, v) plane.

(b) Define $u = x + y$ and $v = x - y$. Sketch the images of the regions S, T and U from question **4.2** in the (u, v) plane.

Ordinary Questions

5.2 The random variables X and Y have joint p.d.f. given by

$$f_{XY}(x,y) = \begin{cases} \frac{1}{2}(x+y)e^{-(x+y)} & \text{for } x, y \geq 0; \\ 0 & \text{elsewhere.} \end{cases}$$

Let $U = X + Y$ and $W = X$.

(a) Find the joint pdf of (U, W) and the marginal p.d.f. of U.

(b) Recognize U as a standard distribution and, using the result of Question **2.5**(a), evaluate $\mathbb{E}[(X + Y)^5]$.

5.3 Let X and Y be standard normal random variables. Show that the joint p.d.f. of (U, V) where $U = X^2$ and $V = X^2 + Y^2$ is given by

$$f_{U,V}(u,v) = \begin{cases} \frac{1}{8\pi} e^{-v^2/2} u^{-1/2}(v-u)^{-1/2} & \text{for } 0 \leq u \leq v \\ 0 & \text{otherwise.} \end{cases}$$

5.4 Let (X, Y) be a random vector with joint p.d.f.

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-(x+y)} & x > y > 0; \\ 0 & \text{otherwise.} \end{cases}$$

(a) If $U = X - Y$ and $V = Y/2$, find the joint p.d.f. of (U, V).

(b) Show that U and V are independent, and recognize their (marginal) distributions as standard distributions.

5.5 Let X and Y be a pair of independent and identically distributed random variables. Let $U = X + Y$ and $V = X - Y$.

(a) Show that U and V are uncorrelated, but give an example to show that U and V are not necessarily independent.

(b) Show that U and V are independent in the special case where X and Y are standard normals.

5.6 Let X and Y be independent random variables with distributions $Ga(\alpha_1, \beta)$ and $Ga(\alpha_2, \beta)$ respectively. Show that the random variables $U = \frac{X}{X+Y}$ and $V = X + Y$ are independent with distributions $Be(\alpha_1, \alpha_2)$ and $Ga(\alpha_1 + \alpha_2, \beta)$ respectively.

5.7 As part of Question **5.6**, we showed that if X and Y are independent random variables with $X \sim Ga(\alpha_1, \beta)$ and $Y \sim Ga(\alpha_2, \beta)$, then $X + Y \sim Ga(\alpha_1 + \alpha_2, \beta)$.

(a) Use induction to show that for $n \geq 2$, if $X_1, X_2, \dots, X_n$ are independent random variables with $X_i \sim Ga(\alpha_i, \beta)$ then

$$\sum_{i=1}^{n} X_i \sim Ga\left(\sum_{i=1}^{n} \alpha_i, \beta\right).$$

(b) Hence show that for $n \geq 1$, if $Z_1, Z_2, \dots, Z_n$ are independent standard normal random variables then

$$\sum_{i=1}^{n} Z_i^2 \sim \chi_n^2.$$

You may use the result of Example 12, which showed that this was true in the case $n = 1$ *(recall that the* χ^2 *distribution is a special case of the Gamma distribution).*

5.8 Let $n \in \mathbb{N}$. The t distribution (often known as Student's t distribution) is the univariate random variable X with p.d.f.

$$f_X(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}},$$

for all $x \in \mathbb{R}$. Here, n is a parameter, known as the number of degrees of freedom.

Let Z be a standard normal random variable and let W be a chi-squared random variable with n degrees of freedom, where Z and W are independent. Show that

$$X = \frac{Z}{\sqrt{W/n}}$$

has the t distribution and n degrees of freedom.

Challenge Questions

5.9 Let $X \sim Exp(\lambda_1)$ and $Y \sim Exp(\lambda_2)$. Show that $U = \min(X, Y)$ has distribution $Exp(\lambda_1 + \lambda_2)$, and that $\mathbb{P}[\min(X, Y) = X] = \frac{\lambda_1}{\lambda_1 + \lambda_2}$. Extend this result, by induction, to handle the minimum of finitely many exponential random variables.

Let $W = \max(X, Y)$. Show that U and $W - U$ are independent.

6 Covariance Matrices and and Multivariate Normal Distributions

Warm-up Questions

6.1 Let

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 4 & -1 \\ -1 & 9 \end{pmatrix}.$$

(a) Calculate $\det(\mathbf{A})$ and find both $\mathbf{A\Sigma}$ and $\mathbf{A\Sigma A}^T$.

(b) Let $\mathbf{A}$ be the vector $(2, 1)$. Show that $\mathbf{A\Sigma A}^T = 21$.

Ordinary Questions

6.2 Let X and Y be independent standard normal random variables.

(a) Write down the covariance matrix of the random vector $\mathbf{X} = (X, Y)^T$.

(b) Let $\mathbf{R}$ be the rotation matrix $\left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$. Show that $\mathbf{RX}$ has the same covariance matrix as $\mathbf{X}$.

6.3 Three (univariate) random variables X, Y and Z have means 3, -4 and 6 respectively and variances 1, 1 and 25 respectively. Further, X and Y are uncorrelated; the correlation coefficient between X and Z is $\frac{1}{5}$ and that between Y and Z is $-\frac{1}{5}$. Let $U = X + Y - Z$ and $W = 2X + Z - 4$.

(a) Write down the mean vector and covariance matrix of $(X, Y, Z)^T$.

(b) Find the mean vector and covariance matrix of $(U, W)^T$.

(c) Evaluate $\mathbb{E}[(2X + Z - 6)^2]$.

6.4 Suppose that the random vector $\mathbf{X} = (X_1, X_2)^T$ follows the bivariate normal distribution with $\mathbb{E}[X_1] = \mathbb{E}[X_2] = 0$, $\mathrm{Var}(X_1) = 1$, $\mathrm{Cov}(X_1, X_2) = 2$ and $\mathrm{Var}(X_2) = 5$.

(a) Calculate the correlation coefficient of X_1 and X_2. Are X_1 and X_2 independent?

(b) Find the mean and the covariance matrices of

$$Y = \begin{pmatrix} 1 & 2 \end{pmatrix} \mathbf{X} \qquad \text{and} \qquad \mathbf{Z} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \mathbf{X}.$$

What are the distributions of Y and $\mathbf{Z}$?

6.5 Let X_1 and X_2 be bivariate normally distributed random variables each with mean 0 and variance 1, and with correlation coefficient ρ.

(a) By integrating out the variable x_2 in the joint p.d.f., verify that the marginal distribution of X_1 is indeed that of a standard univariate normal random variable. *Hint: Use the fact that the integral of a $N(\mu, \sigma^2)$ p.d.f. is equal to* 1.

(b) Show, using the 'usual' formula for the conditional p.d.f. that the conditional p.d.f. of X_2 given $X_1 = x_1$ is $N(\rho x_1, 1-\rho^2)$.

6.6 Let $\mathbf{X} = (X_1, X_2)^T$ have a $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution with mean vector $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$ and covariance matrix $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$. Let

$$\mathbf{A} = \begin{pmatrix} 1 & \frac{-\sigma_1}{\sigma_2} \\ \frac{\sigma_2}{\sigma_1} & 1 \end{pmatrix}.$$

Find the distribution of $\mathbf{Y} = \mathbf{AX}$, and deduce that any bivariate normal random vector can be transformed by a linear transformation into a vector of independent normal random variables.

6.7 The random vector $\mathbf{X} = (X_1, X_2, X_3)^T$ has an $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution where $\boldsymbol{\mu} = (-1, 1, 2)^T$ and

$$\boldsymbol{\Sigma} = \begin{pmatrix} 144 & -30 & 48 \\ -30 & 25 & 10 \\ 48 & 10 & 64 \end{pmatrix}.$$

(a) Find the correlation coefficients between X_1 and X_2, between X_1 and X_3 and between X_2 and X_3.

(b) Let $Y_1 = X_1 + X_3$ and $Y_2 = X_2 - X_1$. Find the distribution of $\mathbf{Y} = (Y_1, Y_2)^T$ and hence find the correlation coefficient between Y_1 and Y_2.

6.8 Recall that an *orthogonal matrix* $\mathbf{R}$ is one for which $\mathbf{R}^{-1} = \mathbf{R}^T$, and recall that if an $n \times n$ matrix $\mathbf{R}$ is orthogonal, $\mathbf{x}$ is an n-dimensional vector, and $\mathbf{y} = \mathbf{Rx}$ then $\sum_{i=1}^n y_i^2 = \mathbf{y} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{x} = \sum_{i=1}^n x_i^2$.
Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be a vector of independent normal random variables with common mean 0 and variance σ^2. Let $\mathbf{R}$ be an orthogonal matrix and let $\mathbf{Y} = \mathbf{RX}$.

(a) Show that $\mathbf{Y}$ is also a vector of independent normal random variables, with common mean 0 and variance σ^2.

(b) Suppose that all the elements in the first row of $\mathbf{R}$ are equal to $\frac{1}{\sqrt{n}}$ (you may assume that an orthogonal matrix exists with this property). Show that $Y_1 = \sqrt{n}\bar{X}$, where $\bar{X}$ is the sample mean of $\mathbf{X}$, and that

$$\sum_{i=2}^n Y_i^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2.$$

(c) Hence, use Question **5.7** to deduce that, if $s^2 = \frac{1}{n-1}\left(\sum_{i=1}^n X_i^2 - n\bar{X}^2\right)$ is the sample variance of $\mathbf{X}$,

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1},$$

and that it is independent of $\bar{X}$.

(d) Let $\mu \in \mathbb{R}$. Deduce that the result of part (c) also holds if the X_i have mean μ (so as they are i.i.d. $N(\mu, \sigma^2)$ random variables).

6.9 Let $Z_1, Z_2, \ldots$ be independent identically distributed normal random variables with mean μ and variance σ^2. We regard n samples of these as a set of data. Write $\bar{Z}$ and s^2 respectively for the sample mean and variance.

Combine **6.8**(d) and **5.8** to show that the statistic

$$X' = \frac{\sqrt{n}(\bar{Z} - \mu)}{s}$$

has the t distribution with $n-1$ degrees of freedom.

Challenge Questions

6.10 Let X be a (univariate) standard normal, with p.d.f. $\phi(x)$, and let $z > 0$. Show that the random variable defined by

$$Y = \begin{cases} X & \text{for } |X| < z \\ -X & \text{for } |X| \geq z \end{cases}$$

has the same distribution as X. Does (X, Y) have a bivariate normal distribution?

7 Likelihood and Maximum Likelihood

Warm-up Questions

7.1 Let $f : \mathbb{R} \to \mathbb{R}$ by $f(\theta) = e^{-\theta^2+4\theta}$.

(a) Find the first derivative of f, and hence identify its turning point(s).

(b) Calculate the value of the second derivative of f at these turning point(s). Hence, deduce if the turning point(s) are local maxima or local minima.

7.2 Let $(a_i)_{i=1}^n$ be a sequence with $a_i \in (0, \infty)$. Show that $\log\left(\prod_{i=1}^n a_i\right) = \sum_{i=1}^n \log a_i$.

Ordinary Questions

7.3 A sample of 3 is obtained from a geometric distribution with unknown parameter θ.

(a) Given this sample, find the likelihood function $L(\theta; 3)$, and state its domain Θ.

(b) Find the maximum likelihood estimator of θ.

7.4 An observation from a $Bi(n, \theta)$ distribution gives the value 4.

(a) Assuming that n is known to be 9 and that θ is unknown, write down the likelihood function of θ, for $\theta \in [0, 1]$, and use a software package of your choice to plot a graph of it.

(b) Find the maximum likelihood estimator of θ.

7.5 A sample (x_1, x_2, x_3) of three observations from a Poisson distribution with parameter λ, where λ is known to be in $\Lambda = \{1, 2, 3\}$, gives the values $x_1 = 4, x_2 = 0, x_3 = 3$. Find the likelihood of each of the possible values of λ, and hence find the maximum likelihood estimate.

7.6 Given the set of i.i.d. samples $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, for some $n > 0$, write down the likelihood function $L(\theta; \mathbf{x})$, in each of the following cases. In each case you should give both the function, in simplified form where possible, and the parameter set Θ.

The data are i.i.d. samples from:

(a) the exponential distribution $Exp(\lambda)$ with $\lambda = 1/\theta$.

(b) the binomial $Bi(m, \theta)$ distribution, where m is known.

(c) the normal $N(\mu, \theta)$, where μ is known.

(d) the gamma distribution $Ga(\theta, 4)$.

(e) the beta distribution $Be(\theta, \theta)$.

7.7 For each of **7.6**(a)-(e), find the log likelihood, and simplify it as much as you can.

7.8 The $IGa(\alpha, \beta)$ distribution is the distribution of $1/U$ if $U \sim Ga(\alpha, \beta)$.

Repeat **7.6** in the case where the data are i.i.d. samples from the inverted gamma distribution $IGa(1, \theta)$. Find and simplify the corresponding log likelihood.

7.9 A set of i.i.d. samples $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ is taken from a geometric distribution with unknown parameter θ, so that the probability function of X_i is

$$p(x_i) = \mathbb{P}[X_i = x_i] = (1-\theta)^{x_i}\theta, \quad x_i = 0, 1, 2, \ldots, \quad 0 < \theta < 1.$$

Find the maximum likelihood estimate of θ based on the above sample.

7.10 Find the maximum likelihood estimate of θ when the data $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ are i.i.d. samples from the binomial $Bi(m, \theta)$ distribution, where m is known, as in question **7.6**/**7.7**(b).

7.11 Find the maximum likelihood estimate of θ when the data $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ are i.i.d. samples from the normal $N(\mu, \theta)$, where μ is known, as in question **7.6**/**7.7**(c). Show that this estimator is unbiased.

7.12 Find the maximum likelihood estimate of $\theta \in (0, \infty)$ when the data $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ are i.i.d. samples from the gamma distribution $Ga(3, \theta)$.

If $\bar{x} = 3$, calculate the maximum likelihood estimate $\hat{\theta}$ and show that it does not depend on the sample size n.

7.13 Suppose that a set of i.i.d samples $\mathbf{x} = (x_1, \ldots, x_n)$ is taken from a negative binomial distribution, so that each X_i has probability function

$$p(x_i) = P(X_i = x_i) = \left(\begin{array}{c} x_i + r - 1 \\ r - 1 \end{array} \right) \theta^r (1-\theta)^{x_i}$$

for some "success" probability θ satisfying $0 \leq \theta \leq 1$, where $x_i = 0, 1, 2, \ldots$ and r is the total number of "successes". If r is known, find the maximum likelihood estimate of θ.

7.14 As in question **7.4**, an observation from a $Bi(n, \theta)$ gives the value $x = 4$.

(a) If θ is known to be $\frac{3}{4}$ but n is unknown, with possible range $n \in \{4, 5, 6, \ldots, \}$, write down a formula for the likelihood function of n and calculate its values for $n = 4, 5, 6$ and 7.

(b) Find the maximum likelihood estimator of n.

7.15 Suppose we have data $\mathbf{x} = (x_1, \ldots, x_n)$, which are i.i.d. samples from a $N(\mu, \sigma^2)$ distribution, where μ is unknown and σ^2 is known.

(a) Find the log-likelihood function of μ.

(b) Show that the maximum likelihood estimator of μ is $\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i$.

(c) Let $k \in (0, \infty)$. Find the k-likelihood region for μ.

7.16 As in question **7.4**(a) an observation from a $Bi(n, \theta)$ distribution with $n = 9$ gives the value $x = 4$.

(a) Find the approximate range of values for which the log likelihood is within 2 of its maximum value. [You may do this either by inspection of a plot, or by using a computer package to solve the inequality numerically.]

(b) A 'traditional' approximate 95% confidence interval here would be of the form $\hat{\theta} \pm 1.96\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$, where $\hat{\theta} = x/n$. Compare your answer to (a) to what this would give.

(c) Repeat this analysis with $n = 90$ and $x = 40$, and comment on your results.

7.17 A set of i.i.d. samples $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ is taken from an inverse Gaussian distribution (known also as the Wald distribution), with p.d.f.

$$f(x_i) = \sqrt{\frac{\theta}{2\pi x_i^3}} \exp\left(-\frac{\theta(x_i - \mu)^2}{2\mu^2 x_i}\right), \quad x_i > 0, \quad \mu, \theta > 0.$$

(a) Assuming that μ is known, find the maximum likelihood estimate of θ based on the above sample.

(b) If both μ and θ are unknown, find the maximum likelihood estimate of (μ, θ) based on the sample given.

Challenge Questions

7.18 A set of i.i.d. samples $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ is taken from a Pareto distribution with unknown parameters α and β so that

$$f(x_i; \boldsymbol{\theta}) = \frac{\alpha\beta^\alpha}{x_i^{\alpha+1}}, \quad x_i \geq \beta > 0, \quad \alpha > 0,$$

where $\boldsymbol{\theta} = (\alpha, \beta)$.

Find the maximum likelihood estimate of $\boldsymbol{\theta}$ based on the above sample.